

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2004, 15:30:06 ; Search time 834 Seconds

(without alignments)  
15905.248 Million cell updates/sec

Title: US-09-856-927-1

Perfect score: 2719  
Sequence: 1 tttaggaacgcacgtgcac.....cattaagtaataagaactt 2719

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3190992 seqs, 2439311697 residues

Total number of hits satisfying chosen parameters: 6381984

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*\n2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*\n3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*\n4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*\n5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*\n6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*\n7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*\n8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*\n9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*\n10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*\n11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*\n12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*\n13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*\n14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*\n15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*\n16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*\n17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*\n18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*\n19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2719	100.0	2719	9	US-09-866-866A-9
2	2700.6	99.3	2718	15	US-10-120-687-60
3	2662.6	97.9	2883	15	US-10-101-510-639
4	2354.6	86.6	2574	9	US-09-981-353-34
5	2352.8	86.5	2418	10	US-09-961-086-2
6	2221.4	81.7	2247	9	US-09-866-866A-16
7	2090.4	76.9	2788	9	US-09-745-763-156
8	2022.2	74.4	2027	16	US-10-405-806-1
9	2010.2	73.9	2053	16	US-10-405-806-12
10	1392.2	51.2	2025	9	US-09-866-866A-13
11	746.4	27.5	795	10	US-09-961-086-7
12	445	16.4	445	13	US-10-083-283-201
13	395.8	14.6	447	9	US-09-960-352-12839
14	364	13.4	427	9	US-09-960-352-12839

C 15	355.4	13.1	467	9	US-09-960-352-9050	Sequence 9050, Ap
C 16	351.6	12.9	414	9	US-09-960-352-5514	Sequence 5514, Ap
C 17	351	12.9	465	9	US-09-983-965-1505	Sequence 1505, Ap
C 18	347.4	12.8	434	15	US-10-313-669-88	Sequence 88, Appl
C 19	347	12.8	456	9	US-09-983-965-1396	Sequence 1396, Ap
C 20	333	12.2	502	9	US-09-866-866A-11	Sequence 11, Appl
C 21	332.2	12.2	410	9	US-09-983-965-1133	Sequence 1133, Ap
C 22	331.2	12.2	408	9	US-09-983-965-1222	Sequence 1222, Ap
C 23	330.4	12.2	377	9	US-09-960-352-4326	Sequence 4326, Ap
C 24	329	12.1	421	9	US-09-983-965-1800	Sequence 1800, Ap
C 25	327.2	12.0	404	9	US-09-983-965-1541	Sequence 1541, Ap
C 26	321.6	11.8	410	9	US-09-983-965-1511	Sequence 1511, Ap
C 27	320.8	11.8	387	9	US-09-960-352-6470	Sequence 6470, Ap
C 28	313.8	11.5	419	9	US-09-983-965-1247	Sequence 1247, Ap
C 29	313.2	11.5	422	9	US-09-960-352-3985	Sequence 3985, Ap
C 30	297.2	10.9	442	9	US-09-983-965-1765	Sequence 1765, Ap
C 31	293.4	10.8	441	9	US-09-983-965-1495	Sequence 1495, Ap
C 32	289.2	10.6	367	9	US-09-983-965-1218	Sequence 1218, Ap
C 33	287.8	10.6	406	9	US-09-960-352-4458	Sequence 4458, Ap
C 34	286.2	10.5	429	9	US-09-983-965-1308	Sequence 1308, Ap
C 35	286	10.5	390	9	US-09-983-965-1431	Sequence 1431, Ap
C 36	285.8	10.5	418	9	US-09-983-965-1351	Sequence 1351, Ap
C 37	285.8	10.4	415	9	US-09-983-965-1136	Sequence 1136, Ap
C 38	282.8	10.3	351	9	US-09-983-965-1435	Sequence 1435, Ap
C 39	279	10.3	414	9	US-09-983-965-1368	Sequence 1368, Ap
C 40	278	10.2	414	9	US-09-983-965-1311	Sequence 1311, Ap
C 41	271.4	10.0	384	9	US-09-983-965-1168	Sequence 1168, Ap
C 42	269.4	9.9	399	9	US-09-983-965-1476	Sequence 1476, Ap
C 43	268.4	9.9	426	9	US-09-983-965-1310	Sequence 1310, Ap
C 44	267.6	9.8	499	9	US-09-833-381-965	Sequence 965, Ap
C 45	261.4	9.6	391	9	US-09-983-965-1108	Sequence 1108, Ap

## ALIGNMENTS

RESULT 1  
US-09-866-866A-9  
Sequence 9, Application US/09866866A  
Patent No. US20020102244A1  
GENERAL INFORMATION:  
APPLICANT: Sorrentino, Brian  
TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
FILE REFERENCE: 1340-1-021CIP2  
CURRENT APPLICATION NUMBER: US/09/866,866A  
CURRENT FILING DATE: 2001-08-30  
PRIOR APPLICATION NUMBER: 09/584,586  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: PCT/US99/11825  
PRIOR FILING DATE: 1999-05-27  
PRIOR APPLICATION NUMBER: 60/086,988  
PRIOR FILING DATE: 1998-05-28  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 9  
LENGTH: 2719  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-866-866A-9  
Query Match 100.0%; Score 2719; DB 9; Length 2719;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 TTAGAAGCAGCCGTCACATCTTGTTGTTTGAAGTGAAGAACTGCTCTTAGA 60  
DB 1 TTAGAAGCAGCCGTCACATCTTGTTGTTTGAAGTGAAGAACTGCTCTTAGA 60  
CY 61 GTTGTGTTGAAGTCCGCTGACATCCCAACATTTACATCTTAATTTGAAGCGC 120  
DB 61 GTTGTGTTGAAGTCCGCTGACATCCCAACATTTACATCTTAATTTGAAGCGC 120

121 TGCCTCGAGCGCAGCATCTCTGAGATCTTGAAGCTTTGGTTAAGCCGAGCTTATTTAA 180  
Db 121 TGCCTCGAGCGCAGCATCTCTGAGATCTTGAAGCTTTGGTTAAGCCGAGCTTATTTAA 180  
Qy 181 GCTGAAAAGATAAAACCTCTCGAGATGCTTTCAGTAAATGTCAGAGTTTATTCAGAGT 240  
Db 181 GCTGAAAAGATAAAACCTCTCGAGATGCTTTCAGTAAATGTCAGAGTTTATTCAGAGT 240  
Qy 241 TCACAGAGAAACACCAATGCTTCCCGCGCAGATTCCTCAATGACCTGAGAGCAATTTACT 300  
Db 241 TCACAGAGAAACACCAATGCTTCCCGCGCAGATTCCTCAATGACCTGAGAGCAATTTACT 300  
Qy 301 GAAGAGAGTGTGTTAAGTTTCAATGATGCTGATGAGTAAATGAGAGAGGCTT 360  
Db 301 GAAGAGAGTGTGTTAAGTTTCAATGATGCTGATGAGTAAATGAGAGAGGCTT 360  
Qy 361 CTACCTTGTGAAAACCAAGTTGAGAAAGAAATATTATCGAATATCAATGAGATGATAA 420  
Db 361 CTACCTTGTGAAAACCAAGTTGAGAAAGAAATATTATCGAATATCAATGAGATGATAA 420  
Qy 421 CCTGTCTCAAGCGCATCTCTGGAGCCACAGGTGAGGCAATCTTGGTTATTTAGATGTC 480  
Db 421 CCTGTCTCAAGCGCATCTCTGGAGCCACAGGTGAGGCAATCTTGGTTATTTAGATGTC 480  
Qy 481 TTAGCTGAGAGAAAGATCCAAAGTGAATTTATCTGAGATGTTCTGATTAATGAGACCG 540  
Db 481 TTAGCTGAGAGAAAGATCCAAAGTGAATTTATCTGAGATGTTCTGATTAATGAGACCG 540  
Qy 541 CGACTGCGCAATTTCAATGATTCAGGTTACGTGATCAAGATGATGTTGTGATGAGC 600  
Db 541 CGACTGCGCAATTTCAATGATTCAGGTTACGTGATCAAGATGATGTTGTGATGAGC 600  
Qy 601 ACTGAGAGGTGAGAGAAACCTTACAGTTCTCAGAGAGCTTCCGCTGCAACAATATG 660  
Db 601 ACTGAGAGGTGAGAGAAACCTTACAGTTCTCAGAGAGCTTCCGCTGCAACAATATG 660  
Qy 661 ACGAATCATGAAAAACGAAAGATTAACAGGCTGATTAAGAGTTAGTCTGATTA 720  
Db 661 ACGAATCATGAAAAACGAAAGATTAACAGGCTGATTAAGAGTTAGTCTGATTA 720  
Qy 721 GTGAGAGATCCAGAGTTGAGAACTCAGTTATCCGTGATGCTGAGAGAGAGAA 780  
Db 721 GTGAGAGATCCAGAGTTGAGAACTCAGTTATCCGTGATGCTGAGAGAGAGAA 780  
Qy 781 AGGACTGATTAAGATGAGAGTTATCACTGATCTTCAATCTTGTCTGATGAGCT 840  
Db 781 AGGACTGATTAAGATGAGAGTTATCACTGATCTTCAATCTTGTCTGATGAGCT 840  
Qy 841 ACAATGAGGTTAGACTCAAGCAGACAAATGCTGCTTTGCTGCTGAAAAGATGCT 900  
Db 841 ACAATGAGGTTAGACTCAAGCAGACAAATGCTGCTTTGCTGCTGAAAAGATGCT 900  
Qy 901 AAGCAGAGAGCAACATCATCTTCTCATTATCAAGCTCGAATTCATCTTCAAGTTG 960  
Db 901 AAGCAGAGAGCAACATCATCTTCTCATTATCAAGCTCGAATTCATCTTCAAGTTG 960  
Qy 961 TTTGATGAGCTCACTTATTTGAGCTCAGAGAACTTAATTTCCAGGAGCTGCTCAGAG 1020  
Db 961 TTTGATGAGCTCACTTATTTGAGCTCAGAGAACTTAATTTCCAGGAGCTGCTCAGAG 1020  
Qy 1021 GCTTGGAGTACTTGAATCAGCTGATTAATCACTGAGAGCTTATTAACCTTGCAGAG 1080  
Db 1021 GCTTGGAGTACTTGAATCAGCTGATTAATCACTGAGAGCTTATTAACCTTGCAGAG 1080  
Qy 1081 TTTCTCTTGGAGATCAATTAATGAGATTCACCTGCTGAGAGCTTAAATTAACCTTGCAGAG 1140  
Db 1081 TTTCTCTTGGAGATCAATTAATGAGATTCACCTGCTGAGAGCTTAAATTAACCTTGCAGAG 1140  
Qy 1141 TTTTAAAGCCAGAGATCATAGAGCTTTCAGAGAGATTAAGCCCATATGAGAAATTA 1200  
Db 1141 TTTTAAAGCCAGAGATCATAGAGCTTTCAGAGAGATTAAGCCCATATGAGAAATTA 1200  
Qy 1201 GGGAGAGTTTATGTCAATCTCTCTCTTCTTCAAAAGAGACAAAGCTGAATTAATCACTT 1260

1201 GGGAGAGTTTATGTCAATCTCTCTCTTCTTCAAAAGAGACAAAGCTGAATTAATCACTT 1260  
Db 1261 TCCGGGGGTGAGAGAGAGAGATCAAGTCTTCAAGAGATGATGATACACACCTCC 1320  
Qy 1261 TCCGGGGGTGAGAGAGAGAGATCAAGTCTTCAAGAGATGATGATACACACCTCC 1320  
Db 1261 TCCGGGGGTGAGAGAGAGAGATCAAGTCTTCAAGAGATGATGATACACACCTCC 1320  
Qy 1321 TTTCTGATCAACTCAGATGAGGTTTCCAGAGCTTCAATTCAAAACCTTGTGGGTAAATCC 1380  
Db 1321 TTTCTGATCAACTCAGATGAGGTTTCCAGAGCTTCAATTCAAAACCTTGTGGGTAAATCC 1380  
Qy 1381 CAGGCTCTATAGCTCAGATCATGTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1440  
Db 1381 CAGGCTCTATAGCTCAGATCATGTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1440  
Qy 1441 TACTTGGCTTAAAAATGATTTCTACTGAAATCCAGAACAGAGCTGGGGTTCTCTTCTTC 1500  
Db 1441 TACTTGGCTTAAAAATGATTTCTACTGAAATCCAGAACAGAGCTGGGGTTCTCTTCTTC 1500  
Qy 1501 CTGAGAGCAACAGAGTTTCAAGAGTGTTCAGAGCTGTTGAGAGCTGTTGTTGTTAGAGAG 1560  
Db 1501 CTGAGAGCAACAGAGTTTCAAGAGTGTTCAGAGCTGTTGAGAGCTGTTGTTGTTAGAGAG 1560  
Qy 1561 AAGCTCTTCAATCATGATTAATCAATCAAGAGATCAAGAGTCAATCTTATTTCTTGA 1620  
Db 1561 AAGCTCTTCAATCATGATTAATCAATCAAGAGATCAAGAGTCAATCTTATTTCTTGA 1620  
Qy 1621 AAATCTGATTCGATTTATTTACCAATGAGAGATTAACCAAGATTAATTTACCTGATA 1680  
Db 1621 AAATCTGATTCGATTTATTTACCAATGAGAGATTAACCAAGATTAATTTACCTGATA 1680  
Qy 1681 GTGTACTCATGTTAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
Db 1681 GTGTACTCATGTTAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
Qy 1741 CTATGATGAGTGTCTTATTTACAGAGTTTCCAGAGTCCAGAGTCCAGAGTCCAGAGT 1800  
Db 1741 CTATGATGAGTGTCTTATTTACAGAGTTTCCAGAGTCCAGAGTCCAGAGTCCAGAGT 1800  
Qy 1801 GTGTGTTCTGAGCAACATCTTCTCATGAGCAATCTGTTTGTGTTATGATGATTTTCA 1860  
Db 1801 GTGTGTTCTGAGCAACATCTTCTCATGAGCAATCTGTTTGTGTTATGATGATTTTCA 1860  
Qy 1861 GGTCTGTGATCAATTCACACATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
Db 1861 GGTCTGTGATCAATTCACACATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
Qy 1921 ATTCCAGATATGATTAAGGCTTTCAGAGATTAATGATTTTGGGACAAAACCTTCTGC 1980  
Db 1921 ATTCCAGATATGATTAAGGCTTTCAGAGATTAATGATTTTGGGACAAAACCTTCTGC 1980  
Qy 1981 CCGAGATCAATGCAACAGAAACATCTTGTATTAATGCAATGATGATGATGATGATGATGAT 2040  
Db 1981 CCGAGATCAATGCAACAGAAACATCTTGTATTAATGCAACATGATGATGATGATGATGAT 2040  
Qy 2041 TATTTGTTAAGAGAGGATTCATCTCACCCTGGGCTTGTGAAAGATCACGAGGC 2100  
Db 2041 TATTTGTTAAGAGAGGATTCATCTCACCCTGGGCTTGTGAAAGATCACGAGGC 2100  
Qy 2101 TTTGCTTGTATGATTTATTTTCTCAGATGCTTACCTGAAATTTGTTATTTCTTAA 2160  
Db 2101 TTTGCTTGTATGATTTATTTTCTCAGATGCTTACCTGAAATTTGTTATTTCTTAA 2160  
Qy 2161 AAATATCTTAAATTTCCCTTATTCAGATGATTTATTCCTCAATTAAGAGAGAGAGAGAG 2220  
Db 2161 AAATATCTTAAATTTCCCTTATTCAGATGATTTATTCCTCAATTAAGAGAGAGAGAGAG 2220  
Qy 2221 TTTGATGAGATTAATCAATCAAGTTTTGTGTTTCTGTTCTTGTCTTGTCTTGTCTTGTCT 2280  
Db 2221 TTTGATGAGATTAATCAATCAAGTTTTGTGTTTCTGTTCTTGTCTTGTCTTGTCTTGTCT 2280  
Qy 2281 TTTGAGAGAGATGTTTAAAGATACATTTTAAATCAACAACTGATTA 2340

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2004, 11:11:46 ; Search time 721 Seconds  
(without alignments)

16020.597 Million cell updates/sec

Title: US-09-856-927-1  
Perfect score: 2719  
Sequence: 1 tttagagacgcacgcgtgcac.....catgaagtaataagaactt 2719

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

- 1: geneseqn1808:\*
- 2: geneseqn1908:\*
- 3: geneseqn20006:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2003as:\*
- 8: geneseqn2003bs:\*
- 9: geneseqn2003cs:\*
- 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2719	100.0	2719	3	AA294760 Human ATP
2	2719	100.0	2719	3	AA27938 ATP-Bind
3	2719	100.0	2719	6	ABA4369 Human BCR
4	2700.6	99.3	2718	7	ACC80605 Human ABC
5	2662.6	97.9	2883	6	AB235528 Human gen
6	2354.6	86.6	2574	4	AA27724 Human tra
7	2354.6	86.6	2574	8	ADA10916 Human CDN
8	2352.8	86.5	2418	2	AA206360 Breast Ca
9	2321.4	81.7	2247	6	ABA4383 Human BCR
10	2090.4	76.9	2788	2	AAV55745 Human sec
11	2090.4	76.9	2788	6	ABQ2071 Human pol
12	2044.2	75.2	2077	4	AAH15008 Human CDN
13	2022.2	74.4	2053	6	ABK49901 CDNA enco
14	2010	73.9	2053	6	ABK49911 CDNA enco
15	1963.2	72.2	1968	9	AD54181 Human bre
16	1963.2	72.2	1968	9	AD54181 Human bre
17	1960	72.1	1998	6	AA142414 Human BCR
18	1955.2	71.9	1998	6	AA142413 Human BCR
19	1932.2	51.2	2025	6	ABA4371 Murine BC
20	746.4	27.5	795	2	AA206365 BCRP frag
21	679.8	25.0	727	4	AAH07859 Human CDN
22	501.4	18.4	594	4	AAH12829 Human CDN
23	461.4	17.0	565	7	AB282925 Toxicolog

24	450.8	16.6	707	3	AA01965	AA01965 Human col
25	445	16.4	445	5	AA293380	AA293380 Lung carc
26	395.8	14.6	447	7	ABX46484	ABX46484 Bovine ES
27	364	13.4	427	7	ABX47674	ABX47674 Bovine ES
28	355.4	13.1	467	7	ABX43885	ABX43885 Bovine ES
29	351.6	12.9	414	7	ABX40349	ABX40349 Bovine ES
30	351	12.9	465	7	ABX51576	ABX51576 Bovine ES
31	347	12.8	456	7	ABX51467	ABX51467 Bovine ES
32	333	12.2	502	3	AA27939	AA27939 ATP-Bind
33	333	12.2	502	6	ABA4370	ABA4370 Mouse BCR
34	332.2	12.2	410	7	ABX51204	ABX51204 Bovine ES
35	331.2	12.2	408	7	ABX51293	ABX51293 Bovine ES
36	330.4	12.2	377	7	ABX39161	ABX39161 Bovine ES
37	329	12.1	421	7	ABX51871	ABX51871 Bovine ES
38	327.2	12.0	404	7	ABX51512	ABX51512 Bovine ES
39	321.6	11.8	410	7	ABX51582	ABX51582 Bovine ES
40	320.8	11.8	387	7	ABX41305	ABX41305 Bovine ES
41	319.8	11.5	419	7	ABX51318	ABX51318 Bovine ES
42	319.2	11.5	422	7	ABX38820	ABX38820 Bovine ES
43	297.2	10.9	442	7	ABX51836	ABX51836 Bovine ES
44	297	10.9	441	7	ABX51566	ABX51566 Bovine ES
45	293.4	10.8	367	7	ABX51289	ABX51289 Bovine ES

#### ALIGNMENTS

RESULT 1	AA294760	standard; CDNA, 2719 BP.
ID	AA294760	
XX		
AC	AA294760;	
XX		
DT	01-AUG-2000	(first entry)
XX		
DE	Human ATP binding cassette ABCG2 CDNA.	
XX		
KM	ABCG2; ATP binding cassette; human; cholesterol; lipid disorder; atherosclerosis; lipid disorder; dyslipidemia; portiasis;	
KM	lupus erythematosus; diagnosis; gene therapy; MPT4;	
KM	multidrug resistance associated protein; Chromosome 4q22-23; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	W0200018912-R2.	
XX		
PD	06-APR-2000.	
XX		
PF	21-SEP-1999;	99WO-EP006991.
XX		
PR	25-SEP-1998;	98US-0101706P.
XX		
PA	(PAB) BAYER AG.	
XX		
PI	Schmitz G, Klucken J;	
XX		
DR	WPI; 2000-293151/25.	
XX		
PT	Adenosine triphosphate binding proteins useful for identifying agents for treating atherosclerosis and other inflammatory disorders.	
XX		
PS	Claim 9; Page 138-139; 154pp; English.	
XX		

The present sequence is that of human ATP binding cassette subfamily C protein ABCG2 CDNA. The CDNA was identified using a differential display method in which monocytes from peripheral blood were subjected to macrophage differentiation and cholesterol loading with acetylated low density lipoproteins and subsequent deacetylation with high density lipoprotein (HDL) to identify cholesterol sensitive genes. The gene maps to chromosome 4q22-23. The invention provides cholesterol-sensitive ABC genes (see AA294734-63). These genes, and polypeptides encoded by them, can be used for diagnostic and therapeutic applications, and for biochemical or cell-based assays to screen for pharmacologically active

modulator compounds useful for the treatment of lipid disorders,  
atherosclerosis or other inflammatory diseases such as psoriasis and  
lupus erythematosus

Sequence 2719 BP: 799 A; 545 C; 564 G; 811 T; 0 U; 0 Other;

Query Match 100.0%; Score 2719; DB 3; Length 2719;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGAGAACGACCGTGACATGCTGGTGGTCTTGTAAAGTGAACCTGCTGTAGA 60  
DB 1 TTGAGAACGACCGTGACATGCTGGTGGTCTTGTAAAGTGAACCTGCTGTAGA 60  
QY 61 GTTTGTTGGAAGGTCCGGGTGACTCATCCCAACATTACATCTTAAATGTTAAAGCGC 120  
DB 61 GTTTGTTGGAAGGTCCGGGTGACTCATCCCAACATTACATCTTAAATGTTAAAGCGC 120  
QY 121 TGCCCTCCGAGCGCAGCATCTCTGAGATCTCTGAGCCTTTGGTTAAGCCGCTCTATTA 180  
DB 121 TGCCCTCCGAGCGCAGCATCTCTGAGATCTCTGAGCCTTTGGTTAAGCCGCTCTATTA 180  
QY 181 GCTGAAAGATTAATAAATCTCTCAGATGCTTTCAGTAATGTCGAAGTTTATATCCAGTG 240  
DB 181 GCTGAAAGATTAATAAATCTCTCAGATGCTTTCAGTAATGTCGAAGTTTATATCCAGTG 240  
QY 241 TCACAGGAAACACCAATGCTTCCCGCGACAGTTTCCATGATCTGAAAGCATTACT 300  
DB 241 TCACAGGAAACACCAATGCTTCCCGCGACAGTTTCCATGATCTGAAAGCATTACT 300  
QY 301 GAAAGAGCTGTGTAAGTTTTCATACATCTGCTATCAGATAAAGTGAAGTGGCTTT 360  
DB 301 GAAAGAGCTGTGTAAGTTTTCATACATCTGCTATCAGATAAAGTGAAGTGGCTTT 360  
QY 361 CTACCTTGTGCAAAAACAGTTGAGAAAGAAATTTATGAAATATCAATGGATATGAAA 420  
DB 361 CTACCTTGTGCAAAAACAGTTGAGAAAGAAATTTATGAAATATCAATGGATATGAAA 420  
QY 421 CCGTGTTCACAGCCATCTCTGGAGCCCAAGGTGAGGCAATCTTGTTATGATGTC 480  
DB 421 CCGTGTTCACAGCCATCTCTGGAGCCCAAGGTGAGGCAATCTTGTTATGATGTC 480  
QY 481 TTAGTGTCAAGAAAGATTCACAGTATTTATCTGAGATGTTCTGATTAATGAGCACCG 540  
DB 481 TTAGTGTCAAGAAAGATTCACAGTATTTATCTGAGATGTTCTGATTAATGAGCACCG 540  
QY 541 CGACCTGCAATTTCAAAATGTAATCAGGTTACGCTGACAAAGATGTTGTGATGGGC 600  
DB 541 CGACCTGCAATTTCAAAATGTAATCAGGTTACGCTGACAAAGATGTTGTGATGGGC 600  
QY 601 ACTGTGACGGTGAAGAAACCTTACAGTTCTTCAGACGCTTCCGCTTGCAACAATATG 660  
DB 601 ACTGTGACGGTGAAGAAACCTTACAGTTCTTCAGACGCTTCCGCTTGCAACAATATG 660  
QY 661 ACGAATCTGAAAAAAGAAAGCAAGGATTAACAGGGTCATTAAGAGTATAGTCTGATATA 720  
DB 661 ACGAATCTGAAAAAAGAAAGCAAGGATTAACAGGGTCATTAAGAGTATAGTCTGATATA 720  
QY 721 GTGGAGACTCCAGAGTTGAACTCAGTTTATCCGCTGCTGTCTGAGAGAAAGAAA 780  
DB 721 GTGGAGACTCCAGAGTTGAACTCAGTTTATCCGCTGCTGTCTGAGAGAAAGAAA 780  
QY 781 AGGACTAGATGAAATGAGAGCTTATCACTGATCTTCCATCTTGTCTGATGAGCT 840  
DB 781 AGGACTAGATGAAATGAGAGCTTATCACTGATCTTCCATCTTGTCTGATGAGCT 840  
QY 841 ACAACTGGCTTAGACTCAGACACAGCAATGCTGTCTTGTCTCTGAAAAAGATGCT 900  
DB 841 ACAACTGGCTTAGACTCAGACACAGCAATGCTGTCTTGTCTCTGAAAAAGATGCT 900  
QY 901 AAGCAGGAGCAACATATCTTCTCATATCAAGCCTCGATATTCATCTTCAAGTTG 960  
DB 901 AAGCAGGAGCAACATATCTTCTCATATCAAGCCTCGATATTCATCTTCAAGTTG 960

QY 961 TTGATAGCTCACCTTATTTGGCTCAGAGAACTATTTCCAGGCGCTGCTCAGAG 1020  
DB 961 TTGATAGCTCACCTTATTTGGCTCAGAGAACTATTTCCAGGCGCTGCTCAGAG 1020  
QY 1021 GCGTTGGGATCTTTGAATCAGCTGTGTATCACTGTAGGGCTTATATACCCTGCAAGC 1080  
DB 1021 GCGTTGGGATCTTTGAATCAGCTGTGTATCACTGTAGGGCTTATATACCCTGCAAGC 1080  
QY 1081 TTCTTCTTGACATCATTAATGAGATTTCCACTGCTGTGCAATTAAACAGAAAGAAC 1140  
DB 1081 TTCTTCTTGACATCATTAATGAGATTTCCACTGCTGTGCAATTAAACAGAAAGAAC 1140  
QY 1141 TTTAAAGCCACAGATCATAGAGCTTCCAGCAGATTAAGCCATCATTAAGAAAATTA 1200  
DB 1141 TTTAAAGCCACAGATCATAGAGCTTCCAGCAGATTAAGCCATCATTAAGAAAATTA 1200  
QY 1201 GCGAGATTTATGCACTCCCTCTTACAAAGAGACAAAAGCTGAATTCATCACTT 1260  
DB 1201 GCGAGATTTATGCACTCCCTCTTACAAAGAGACAAAAGCTGAATTCATCACTT 1260  
QY 1261 TCCGCGGCTGAGAGAAAGAAAGATCAAGTCTTCAAGAGATCACTACACACCTCC 1320  
DB 1261 TCCGCGGCTGAGAGAAAGAAAGATCAAGTCTTCAAGAGATCACTACACACCTCC 1320  
QY 1321 TTCTGTATCACTGAGATGGGTTTCCAGAGCTTCAATTAATACTGTGGGTAAATCC 1380  
DB 1321 TTCTGTATCACTGAGATGGGTTTCCAGAGCTTCAATTAATACTGTGGGTAAATCC 1380  
QY 1381 CAGGCTCTATAGCTCAGATCATGTCACAGTGTATGGAACCTGTTATAGTGCAT 1440  
DB 1381 CAGGCTCTATAGCTCAGATCATGTCACAGTGTATGGAACCTGTTATAGTGCAT 1440  
QY 1441 TACTTTGGCTTAAAAATGATTTCTACTGGAATCCAGAAACAGCTGGGGTTCTCTTC 1500  
DB 1441 TACTTTGGCTTAAAAATGATTTCTACTGGAATCCAGAAACAGCTGGGGTTCTCTTC 1500  
QY 1501 CTGACAGCAACACAGTGTTCAGAGTGTTCAGCCGTGGAACCTTTGTGTAGAGAG 1560  
DB 1501 CTGACAGCAACACAGTGTTCAGAGTGTTCAGCCGTGGAACCTTTGTGTAGAGAG 1560  
QY 1561 AAGCTTTATCATGAAATACATCAGCGGATCTACAGAGTCAATTTATCTCTGGA 1620  
DB 1561 AAGCTTTATCATGAAATACATCAGCGGATCTACAGAGTCAATTTATCTCTGGA 1620  
QY 1621 AAACGTATATCTGATTTATTAACCATAGAGATTAACAAAGTATTAATTAACCTGATA 1680  
DB 1621 AAACGTATATCTGATTTATTAACCATAGAGATTAACAAAGTATTAATTAACCTGATA 1680  
QY 1681 GTGTACTTCATGTTAGGATTTGAAGCCAAAGCAGATGCTTCTGTGTATGATTTACC 1740  
DB 1681 GTGTACTTCATGTTAGGATTTGAAGCCAAAGCAGATGCTTCTGTGTATGATTTACC 1740  
QY 1741 CTTATGATGCTGCTTATCAAGCAGTTCATGCACTGCGCATAGCAGAGT 1800  
DB 1741 CTTATGATGCTGCTTATCAAGCAGTTCATGCACTGCGCATAGCAGAGT 1800  
QY 1801 GTGGTTTCTGAGCAACCTTCTCATGACATCTGTTTGTGTTATGATGATTTTCTA 1860  
DB 1801 GTGGTTTCTGAGCAACCTTCTCATGACATCTGTTTGTGTTATGATGATTTTCTA 1860  
QY 1861 GGTCTGTGGTCAATCTCAACAACATTTGATCTGAGCTGTCAAGTCACTCAAG 1920  
DB 1861 GGTCTGTGGTCAATCTCAACAACATTTGATCTGAGCTGTCAAGTCACTCAAG 1920  
QY 1921 ATTCCAGATATGATTTACGGCTTTGACAGATATGAAATTTTGGACAAAATTTCTGC 1980  
DB 1921 ATTCCAGATATGATTTACGGCTTTGACAGATATGAAATTTTGGACAAAATTTCTGC 1980  
QY 1981 CGAGACTCAATGCAACAGAAACAAATCTTGTAACTATAGCAACATGTAAGCCAGAGAA 2040  
DB 1981 CGAGACTCAATGCAACAGAAACAAATCTTGTAACTATGCAACATGTAAGCCAGAGAA 2040



361 CTACCTTGTGAAAAACAGTTGAGAAAGAAATATTATGAAATATCATGAGATCATGAAA 420  
365 CTACCTTGTGAAAAACAGTTGAGAAAGAAATATTATGAAATATCATGAGATCATGAAA 454  
421 COTGCTTCACAGCCATCTCTGGAGCCCAAGTGGAGGCAATCTTCTGTTTATGATGTC 480  
445 CCGTCTTCACAGCCATCTCTGGAGCCCAAGTGGAGGCAATCTTCTGTTTATGATGTC 514  
481 TTAGCTGAGAGAAAGATCCAAAGTGAATTAATCTGAGATGTTCTGATTAATGAGACCG 540  
515 TTAGCTGAGAGAAAGATCCAAAGTGAATTAATCTGAGATGTTCTGATTAATGAGACCG 574  
541 CGACTGCGCAATTTCAAAATGTAATTCAGATTACGTGATCAAGATGATGTTGATGAGGC 600  
575 CGACTGCGCAATTTCAAAATGTAATTCAGATTACGTGATCAAGATGATGTTGATGAGGC 634  
601 ACTCTGACGTGAGAGAAAACTTACAGTTCTCAGACAGCTCTTGGCTTGCACAACTATG 660  
635 ACTCTGACGTGAGAGAAAACTTACAGTTCTCAGACAGCTCTTGGCTTGCACAACTATG 694  
661 ACGAATCTGAAAAAAGCAAGGATTAACAGGCTCATTTGAAGGTTAGTCTGAGATPAA 720  
695 ACGAATCTGAAAAAAGCAAGGATTAACAGGCTCATTTGAAGGTTAGTCTGAGATPAA 754  
721 GTGGAGAGCTCCAGAGTTGGAATCAGTTTATCCGTGCTGATGCTGAGAGAGAAAA 780  
755 GTGGAGAGCTCCAGAGTTGGAATCAGTTTATCCGTGCTGATGCTGAGAGAGAAAA 814  
781 AGGATTAATGAGATGAGATGAGATTCATGATCTTCATCTTGTCTGATGAGCT 840  
815 AGGATTAATGAGATGAGATGAGATTCATGATCTTCATCTTGTCTGATGAGCT 874  
841 ACAATGCTTAGATAGATCAAGACAGCAAAATGCTGCTTGTCTGCTGAAAAAGATGCT 900  
875 ACAATGCTTAGATAGATCAAGACAGCAAAATGCTGCTTGTCTGCTGAAAAAGATGCT 934  
901 AAGCAGGAGCAACATCATCTTCTCATTCATCAGCTCGATATTCATCTTCAAGTTG 960  
935 AAGCAGGAGCAACATCATCTTCTCATTCATCAGCTCGATATTCATCTTCAAGTTG 994  
961 TTGATTAACCTCACCCTTATGCTGCTCAGAGAACTTAATGTTCCACGGCTTGTCAAGAG 1020  
995 TTGATTAACCTCACCCTTATGCTGCTCAGAGAACTTAATGTTCCACGGCTTGTCAAGAG 1054  
1021 GCTTGGGATCTTGAATCAGCTGCTTATCATCTGAGGCTTAAATTAACCTGCAAGC 1080  
1055 GCTTGGGATCTTGAATCAGCTGCTTATCATCTGAGGCTTAAATTAACCTGCAAGC 1114  
1081 TTCTTCTTGACATATTAATGAGATTCACCTGCTGCAATTAACAGAGAAAGAGC 1140  
1115 TTCTTCTTGACATATTAATGAGATTCACCTGCTGCAATTAACAGAGAAAGAGC 1174  
1141 TTAAAGCACAAGATCATAGAGCTTCAAGCAGATAGCACTCATAGAAAAATTA 1200  
1175 TTAAAGCACAAGATCATAGAGCTTCAAGCAGATAGCACTCATAGAAAAATTA 1234  
1201 GCGAGATTTATGTAATCTCTCTCTTCAAAAGAGACAAAGCTGAATTCATCAACTT 1260  
1235 GCGAGATTTATGTAATCTCTCTCTTCAAAAGAGACAAAGCTGAATTCATCAACTT 1294  
1261 TCCGGGGGTGAGAGAAAGAGATCAAGCTTCAAGAGATCAAGCTCAACCACTCC 1320  
1295 TCCGGGGGTGAGAGAAAGAGATCAAGCTTCAAGAGATCAAGCTCAACCACTCC 1354  
1321 TTCTGTCACTCACTCAAGTGGCTTCCAGAGCTTATTCAAAAAATCTGCTGGATPAACTCC 1380  
1355 TTCTGTCACTCACTCAAGTGGCTTCCAGAGCTTATTCAAAAAATCTGCTGGATPAACTCC 1414  
1381 CAGGCTCTATAGCTCAATCATCTTGTCAAGCTGATGAGACTGTTATAGTGGCATT 1440  
1415 CAGGCTCTATAGCTCAATCATCTTGTCAAGCTGATGAGACTGTTATAGTGGCATT 1474

1441 TACTTGGGCTTAAAAATGATTTCTACTGAAATCCAGAAACAGAGCTGGGCTTCTCTTTC 1500  
1475 TACTTGGGCTTAAAAATGATTTCTACTGAAATCCAGAAACAGAGCTGGGCTTCTCTTTC 1534  
1501 CTGAGAGCAACAGAGTTTCCAGAGCTGTTTACGCGTGGAACTCTTGTGTGATAGAG 1560  
1535 CTGAGAGCAACAGAGTTTCCAGAGCTGTTTACGCGTGGAACTCTTGTGTGATAGAG 1594  
1561 AAGCTTTCATACATGATATCAACAGCGATATCAAGAGTGCATATTTCTTGGAA 1620  
1595 AAGCTTTCATACATGATATCAACAGCGATATCAAGAGTGCATATTTCTTGGAA 1654  
1621 AAGCTTTCATACATGATATTAACCAAGAGTATCAAGAGTATATTAATTTCTGATA 1680  
1655 AAGCTTTCATACATGATATTAACCAAGAGTATCAAGAGTATATTAATTTCTGATA 1714  
1681 GTGTACTTCATGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 1740  
1715 GTGTACTTCATGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 1774  
1741 CTATATAGTGTGCTTATTCAGCAGTTTCATGAGCACTGAGCATATGACAGGTCAGAT 1800  
1775 CTATATAGTGTGCTTATTCAGCAGTTTCATGAGCACTGAGCATATGACAGGTCAGAT 1834  
1801 GTGTACTTCATGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 1860  
1835 GTGTACTTCATGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 1894  
1861 GGTCTGTGTGATGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 1920  
1895 GGTCTGTGTGATGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 1954  
1921 ATTCCAGATATGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 1980  
1955 ATTCCAGATATGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 2014  
1981 CAGAGCTCATGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 2040  
2015 CAGAGCTCATGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 2074  
2041 TATTGTTAAGAGGAGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 2100  
2075 TATTGTTAAGAGGAGATGATGATGAGTGAAGCAAGGAGATGCTTCTGTTATGATGTTAAC 2134  
2101 TTGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
2135 TTGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2194  
2161 AATATTTCTTAATTTCCCTTAATTCAGTATGATGATGATGATGATGATGATGATGATGAT 2220  
2195 AATATTTCTTAATTTCCCTTAATTCAGTATGATGATGATGATGATGATGATGATGATGAT 2254  
2221 TTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
2255 TTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2314  
2281 TTGACAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
2315 TTGACAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2374  
2341 AACATGAAAGAACCCAGACATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2372  
2375 AACATGAAAGAACCCAGACATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2406

RESULT 2  
US-09-245-808-7  
; Sequence 7, Application US/09245808  
; Patent No. 6313277  
; GENERAL INFORMATION:  
; APPLICANT: Doyle, L. Austin  
; APPLICANT: Abruzzo, Lynne V.  
; APPLICANT: Ross, Douglas D.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2004, 13:01:06 ; Search time 4551 Seconds

(without alignments)  
17841.193 Million cell updates/sec

Title: US-09-856-927-1

Perfect score: 2719

Sequence: 1 tttaggaacgacacgtgcac.....catlaagtaataagaactt 2719

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
EST:  
1: em\_estda:\*  
2: em\_esthm:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_escov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gse\_hum:\*  
18: em\_gse\_inv:\*  
19: em\_gse\_pin:\*  
20: em\_gse\_vrt:\*  
21: em\_gse\_fun:\*  
22: em\_gse\_mam:\*  
23: em\_gse\_mus:\*  
24: em\_gse\_pro:\*  
25: em\_gse\_rtd:\*  
26: em\_gse\_phg:\*  
27: em\_gse\_vr1:\*  
28: gb\_gsal:\*  
29: gb\_gsal2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	982.6	36.1	1345	11	U66681 Homo sapien
2	970.2	35.7	2869	11	AK030831 Mus muscu
3	970.2	35.7	3001	11	AK051880 Mus muscu
4	913	33.6	1034	12	BMS44411 AGENCOURT

5	885.4	32.6	925	13	BX370114	BX370114
6	859.8	31.6	1201	13 <td>BX382704 <td>BX382704</td> </td>	BX382704 <td>BX382704</td>	BX382704
7	847.2	31.2	907	13 <td>BX329115 <td>BX329115</td> </td>	BX329115 <td>BX329115</td>	BX329115
8	821.2	30.2	888	13 <td>BX454801 <td>BX454801</td> </td>	BX454801 <td>BX454801</td>	BX454801
9	816.4	30.0	883	9	AU118354 <td>AU118354</td>	AU118354
10	782.6	28.8	1201	13 <td>BX382703 <td>BX382703</td> </td>	BX382703 <td>BX382703</td>	BX382703
11	773	28.4	839	12 <td>BG53081 <td>BG53081</td> </td>	BG53081 <td>BG53081</td>	BG53081
12	753.8	27.7	821	12 <td>BI918550 <td>BI918550</td> </td>	BI918550 <td>BI918550</td>	BI918550
13	725.6	26.7	794	10 <td>AM155343 <td>AM155343</td> </td>	AM155343 <td>AM155343</td>	AM155343
14	714	26.3	740	12 <td>BG537904 <td>BG537904</td> </td>	BG537904 <td>BG537904</td>	BG537904
15	679.8	25.0	727	9	AU137863 <td>AU137863</td>	AU137863
16	679.4	25.0	836	14 <td>CF593368 <td>CF593368</td> </td>	CF593368 <td>CF593368</td>	CF593368
17	678.6	25.0	961	12 <td>BG386140 <td>BG386140</td> </td>	BG386140 <td>BG386140</td>	BG386140
18	668.6	24.6	739	11 <td>BC011032 <td>BC011032</td> </td>	BC011032 <td>BC011032</td>	BC011032
19	666.2	24.5	708	14 <td>CD355766 <td>CD355766</td> </td>	CD355766 <td>CD355766</td>	CD355766
20	641.2	23.6	666	9 <td>AL135203 <td>AL135203</td> </td>	AL135203 <td>AL135203</td>	AL135203
21	637.4	23.4	697	9 <td>AV728650 <td>AV728650</td> </td>	AV728650 <td>AV728650</td>	AV728650
22	632	23.2	879	12 <td>BM423310 <td>BM423310</td> </td>	BM423310 <td>BM423310</td>	BM423310
23	624.8	23.0	631	12 <td>BI497116 <td>BI497116</td> </td>	BI497116 <td>BI497116</td>	BI497116
24	614.4	22.6	950	13 <td>BX370198 <td>BX370198</td> </td>	BX370198 <td>BX370198</td>	BX370198
25	611.2	22.5	620	9 <td>AL603604 <td>AL603604</td> </td>	AL603604 <td>AL603604</td>	AL603604
26	608.4	22.4	787	13 <td>BX418619 <td>BX418619</td> </td>	BX418619 <td>BX418619</td>	BX418619
27	607	22.3	671	12 <td>BI086640 <td>BI086640</td> </td>	BI086640 <td>BI086640</td>	BI086640
28	592.6	21.8	664	12 <td>BG436056 <td>BG436056</td> </td>	BG436056 <td>BG436056</td>	BG436056
29	586.6	21.6	644	14 <td>CB552639 <td>CB552639</td> </td>	CB552639 <td>CB552639</td>	CB552639
30	584	21.5	954	13 <td>BX370199 <td>BX370199</td> </td>	BX370199 <td>BX370199</td>	BX370199
31	567.4	20.9	588	10 <td>BE048216 <td>BE048216</td> </td>	BE048216 <td>BE048216</td>	BE048216
32	562.4	20.7	607	12 <td>BI553320 <td>BI553320</td> </td>	BI553320 <td>BI553320</td>	BI553320
33	558.2	20.5	712	12 <td>BG689713 <td>BG689713</td> </td>	BG689713 <td>BG689713</td>	BG689713
34	558	20.5	873	10 <td>BF207763 <td>BF207763</td> </td>	BF207763 <td>BF207763</td>	BF207763
35	529.8	19.5	907	13 <td>BX409712 <td>BX409712</td> </td>	BX409712 <td>BX409712</td>	BX409712
36	518.6	19.1	653	10 <td>BF674699 <td>BF674699</td> </td>	BF674699 <td>BF674699</td>	BF674699
37	506.8	18.6	672	14 <td>CB424319 <td>CB424319</td> </td>	CB424319 <td>CB424319</td>	CB424319
38	501.4	18.4	594	9 <td>AU157347 <td>AU157347</td> </td>	AU157347 <td>AU157347</td>	AU157347
39	500.4	18.4	534	9 <td>AU144940 <td>AU144940</td> </td>	AU144940 <td>AU144940</td>	AU144940
40	495	18.2	586	12 <td>BG705688 <td>BG705688</td> </td>	BG705688 <td>BG705688</td>	BG705688
41	493.2	18.1	498	9 <td>AI872416 <td>AI872416</td> </td>	AI872416 <td>AI872416</td>	AI872416
42	492.4	18.1	722	14 <td>CB959914 <td>CB959914</td> </td>	CB959914 <td>CB959914</td>	CB959914
43	492	18.1	492	10 <td>AM270740 <td>AM270740</td> </td>	AM270740 <td>AM270740</td>	AM270740
44	478	17.6	489	9 <td>AA933960 <td>AA933960</td> </td>	AA933960 <td>AA933960</td>	AA933960
45	473.4	17.4	475	10 <td>BE672952 <td>BE672952</td> </td>	BE672952 <td>BE672952</td>	BE672952

## ALIGNMENTS

RESULT 1  
LOCUS U66681 1345 bp mRNA linear HTC 23-JUL-2001  
DEFINITION Homo sapiens clone EST157461 mRNA sequence.  
ACCESSION U66681  
VERSION U66681.1 GI:1906566  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1345)  
AUTHORS Allikmets,R., Gerrard,B., Hutchinson,A. and Dean,M.  
TITLE Characterization of the human ABC superfamily: isolation and mapping of 21 new genes using the expressed sequence tags database Hum. Mol. Genet. 5 (10), 1649-1655 (1996)  
JOURNAL JOURNAL  
MEDLINE 97049974  
PUBMED 8894702  
REFERENCE 2 (bases 1 to 1345)  
AUTHORS Allikmets,R., Gerrard,B. and Dean,M.  
TITLE Direct Submission  
JOURNAL Submitted (12-AUG-1996) Human Genetics Section, National Cancer Institute, NCI-PCRDC, Frederick, MD 21702, USA  
FEATURES  
source  
1..1345  
/organism="Homo sapiens"  
/mol\_type="mRNA"







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 18, 2004, 11:16:26 ; Search time 6989 Seconds

(without alignments)  
16862.170 Million cell updates/sec

Title: US-09-856-927-1  
Sequence: 1 tttaggaacgcacgcgtgcac.....cattaagatcatagacct 2719

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_ncg.\*  
3: gb\_in.\*  
4: gb\_cm.\*  
5: gb\_cv.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pt.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_pa.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vl.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rtd.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_ey.\*  
39: em\_hgtgo\_hum.\*  
40: em\_hgtgo\_mus.\*  
41: em\_hgtgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2719	100.0	2719	6 BD344217	BD344217 ATP-bind
2	2719	100.0	2719	6 AX322795	AX322795 Sequence
3	2719	100.0	2719	6 AF103796	AF103796 Homo sapi
4	2354.6	86.6	2574	6 AX061220	AX061220 Sequence
5	2354.4	86.6	2418	6 AF098951	AF098951 Homo sapi
6	2352.8	86.5	2418	6 AF177881	AF177881 Sequence
7	2352.8	86.5	2418	6 BD124164	BD124164 Breast ca
8	2332.4	82.1	2547	6 BC021281	BC021281 Homo sapi
9	2221.4	81.7	2247	6 AX322812	AX322812 Sequence
10	2221.4	81.7	2247	6 AF017168	AF017168 Homo sapi
11	2212	81.4	2212	6 AY89766	AY89766 Homo sapi
12	2090.4	76.9	2788	6 BD192221	BD192221 Secreted
13	2078.6	76.4	2085	6 AF463519	AF463519 Homo sapi
14	2044.2	75.2	2077	6 AX878056	AX878056 Sequence
15	2044.2	75.2	2077	6 BD157000	BD157000 Primer fo
16	2044.2	75.2	2077	6 AK002040	AK002040 Homo sapi
17	2022.2	74.4	2027	6 BD167779	BD167779 Drug-tole
18	2022.2	74.4	2027	6 AB051855	AB051855 Homo sapi
19	2010	73.9	2053	6 BD167789	BD167789 Drug-tole
20	1963.2	72.2	1968	6 BD133828	BD133828 Variant B
21	1963.2	72.2	1968	6 AB056867	AB056867 Homo sapi
22	1960	72.1	1998	6 BD133830	BD133830 Variant B
23	1955.2	71.9	1998	6 BD133829	BD133829 Variant B
24	1868.6	62.0	1690	6 AY288307	AY288307 Homo sapi
25	1506.4	55.4	2112	4 SSC420927	AJ420927 Sus scrofa
26	1431.8	52.7	1469	9 AF093771	AF093771 Homo sapi
27	1408	51.8	2175	10 AY089997	AY089997 Rattus no
28	1403.2	51.5	2139	10 AY089996	AY089996 Rattus no
29	1400	51.5	2130	10 AY089998	AY089998 Rattus no
30	1398.8	51.4	2083	10 AB105817	AB105817 Rattus no
31	1393	51.2	2520	10 BC053730	BC053730 Mus muscu
32	1392.2	51.2	2025	6 AX322799	AX322799 Sequence
33	1392.2	50.1	2025	10 AF140218	AF140218 Mus muscu
34	1363	50.1	1974	10 AB094089	AB094089 Rattus no
35	1303.6	47.9	176392	2 AC135954	AC135954 Papio anu
36	1240	45.6	187712	9 AC027088	AC027088 Homo sapi
37	1103.6	40.6	1144	9 AF093772	AF093772 Homo sapi
38	971.8	35.7	2277	10 AF324242	AF324242 Mus muscu
39	746.4	27.5	795	6 AR177886	AR177886 Sequence
40	746.4	27.5	795	6 BD124169	BD124169 Breast ca
41	682	25.1	181576	9 AC084732	AC084732 Homo sapi
42	679.8	25.0	727	6 AX869789	AX869789 Sequence
43	679.8	25.0	727	6 BD149851	BD149851 Primer fo
44	574.6	21.1	231679	2 AC123319	AC123319 Rattus no
45	574.6	21.1	255881	2 AC099252	AC099252 Rattus no

#### ALIGNMENTS

RESULT 1  
BD344217 2719 bp DNA linear PAT 17-UTL-2003  
LOCUS BD344217  
DEFINITION ATP-binding cassette genes and proteins for diagnosis and remedy of  
lipid disorders and inflammatory diseases.  
ACCESSION BD344217  
VERSION BD344217.1 GI:33043587  
KEYWORDS JP 2002525111-A/27.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 2719)  
Schmitz, G. and Klucken, J.  
TITLE ATP-binding cassette genes and proteins for diagnosis and remedy of

